UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/901,031	07/10/2001	Dan Chazan	CHAZAN=1A	6360

1444 7590 06/08/2005

BROWDY AND NEIMARK, P.L.L.C.
624 NINTH STREET, NW
SUITE 300
WASHINGTON, DC 20001-5303

EXAMINER	
JACKSON, JAKIEDA R	

ART UNIT	PAPER NUMBER
2655	

DATE MAILED: 06/08/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/901,031	CHAZAN ET AL.	
	Examiner	Art Unit	
	Jakieda R Jackson	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *30 December 2004*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-75* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-75* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *12/30/04*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Response to Amendment

1. In response to the Office Action mailed September 30, 2004, applicant submitted an amendment filed on December 30, 2004, in which the applicant traversed and requested reconsideration with respect to **independent claims 1, 16, 26, 41, 51 and 66**.

Response to Arguments

2. Applicant's argue regarding claim 1 that Itoh's speech synthesizer is based on concatenating time-domain waveforms, meanwhile, claim 1 is directed to feature-domain concatenation. Applicant's argue that Itoh explicitly teaches away from feature domain concatenation approaches. Applicant's also argue that Itoh's clustering of LPC parameter vectors, describes clustering as gathering together similar phonemes from differenct parts of the input speech database. Applicant's argue that it has nothing to do with generating an ordered, concatenated output series for speech synthesis. Also, regarding claim 1, applicant's argue that there is simply no mention or suggestion of "complex line spectrum", which is defined as "the sequence of respective sine-wave amplitudes. Phrases and frequencies in a sinusoidal speech representation".

Regarding claim 16, applicant's argue that itoh's phoneme waveforms are time-domain entities. Itoh teaches away from parameter-domain concatenation.

Independent claims 26 and 41 recite devices for speech synthesis while independent claims 51 and 66 recite computer software products. The claimed devices and products operate on principles similar to the methods of claim 1 and 16 and

therefore, applicant's argue that they are patentable for the same reasons as stated above.

Applicant's arguments, see pages 25-33, filed December 30, 2004, with respect to the rejection(s)of claim(s) 1, 16, 26, 41, 51 and 66 under 102(b) have been fully considered and are persuasive. Therefore, the rejection has been withdrawn. However, upon further consideration, a new ground(s) of rejection is made in view of Aso.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. **Claims 1-8, 10, 13-22, 26-33, 35, 38-47, 51-58, and 63-72** are rejected under 35 U.S.C. 103(a) as being unpatentable over Itoh (U.S. Patent No. 5,740,320) in view of Aso (USPN 5,485,543).

Regarding **claims 1, 26 and 51**, Itoh discloses the method, device and computer software product for speech synthesis, comprising:

providing a segment inventory comprising, for a plurality of speech segments, respective sequences of feature vectors (phoneme string; column 6, lines 9-17), by

estimating spectral envelopes (spectrum envelope) of input speech signals corresponding to the speech segments in a succession of time intervals during each of the speech segments, and integrating the spectral envelopes over a plurality of window functions (analysis window shifting) in a frequency domain so as to determine vector elements of the feature vectors (column 5, lines 19-35 with column 5, line 64 – column 6, line 3);

receiving phonetic (phoneme) and prosodic information (prosodic information) indicative of an output speech signal to be generated (column 6, lines 7-18); and

selecting the sequences of feature vectors (phoneme string) from the inventory responsive to the phonetic (phoneme) and prosodic information (prosodic information; column 9, lines 14-23), but lacks feature domain concatenation.

Aso discloses a method for speech synthesis, comprising:

processing the selected sequences of feature vectors (figures 1 and 6 with train of sample points; column 4, lines 1-8) so as to generate a concatenated output series of feature vectors (converting into mel cepstrum; column 2, lines 55-67 with column 3, lines 15-22 and lines 53-67);

computing a series of complex line spectra of the output signal from the series of the feature vectors (train sample points; column 4, lines 1-8 and column 2, lines 55-67 with column 3, lines 15-22); and

transforming the complex line spectra to a time domain speech signal for output (column 6, lines 24-31), obtaining a synthesized speech of higher quality.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Itoh's method, device and computer software product wherein it comprises feature domain concatenation, to obtain high-quality synthesized speech with low complexity in a more practical manner (column 2, lines 21-32).

Regarding **claims 2, 27 and 52**, Itoh discloses a method, device and computer software product wherein providing the segment inventory comprises providing segment information comprising respective phonetic identifiers of the segments (label indicating combination of phoneme; column 5, lines 1-2), and wherein selecting the sequences of feature vectors comprises finding the segments whose phonetic identifiers are close to the received phonetic information (closest to; column 5, lines 59-64).

Regarding **claims 3, 28 and 53**, Itoh discloses a method, device and computer software product method wherein the segments comprise lefemes (first/third phoneme), and wherein the phonetic identifiers comprise lefeme labels (label indicating combination; column 5, lines 1-17).

Regarding claims **4, 29 and 54**, Itoh discloses a method, device and computer software product wherein the segment information further comprises one or more prosodic parameters with respect to each of the segments (prosodic information), and wherein selecting the sequences of feature vectors (phoneme string; column 6, lines 7-31) comprises finding the segments whose one or more prosodic parameters are close

to the received prosodic information (parameters are close to the centroids; column 7, lines 41-44).

Regarding **claims 5, 30 and 55**, Itoh discloses a method, device and computer software product wherein the one or more prosodic parameters are selected from a group of parameters consisting of a duration (duration), an energy level (power variation) and a pitch (pitch) of each of the segments (column 6, lines 7-31 with column 9, lines 20-23).

Regarding **claims 6, 31 and 56**, Itoh discloses a method, device and computer software product wherein the feature vectors comprise auxiliary vector elements (clusters represent a spectrum domain) indicative of further features of the speech segments, in addition to the elements determined by integrating the spectral envelopes (spectrum envelope) of the input speech signals (column 5, lines 47-64).

Regarding **claims 7, 32 and 57**, Itoh discloses a method, device and computer software product wherein the auxiliary vector elements comprise voicing vector elements indicative of a degree of voicing of frames of the corresponding speech segments (degree of phoneme; column 9, lines 53-67), and wherein computing the complex line spectra comprises reconstructing the output speech signal with the degree of voicing indicated by the voicing vector elements (degree of phoneme; column 9, lines 53-67).

Regarding **claims 8, 33 and 58**, Itoh discloses a method, device and computer software product wherein receiving the prosodic information comprises receiving pitch values (pitch), and wherein reconstructing the output speech signal comprises

adjusting a frequency spectrum (spectrum) of the output speech signal responsive to the pitch values (column 5, line 64 –column 6, line 31 and column 9, lines 14-23).

Regarding **claims 10, 35 and 60**, Itoh discloses a method, device and computer software product wherein concatenating the selected sequences of feature vectors (phoneme string) comprises adjusting the feature vectors responsive to the prosodic information (prosodic information; column 6, lines 7-31 and column 9, lines 14-23).

Regarding **claims 13, 38 and 63**, Itoh discloses a method, device and computer software product wherein the prosodic information comprises respective energy levels (power pattern) of the segments to be incorporated in the output speech signal, and wherein adjusting the feature vectors comprises altering one or more of the vector elements so as to adjust the energy levels of one or more of the segments (set desired power pattern; column 9, lines 14-23 with column 6, lines 18-28).

Regarding **claims 14, 39 and 64**, Itoh discloses a method, device and computer software product wherein processing the selected sequences (phoneme string) comprises adjusting the vector elements so as to provide a smooth transition between the segments in the time domain signal (phoneme segments smoothly continue; column 6, lines 7-31).

Regarding **claims 15, 40 and 65**, Itoh discloses a method, device and computer software product wherein the vector elements comprise Mel Frequency Cepstral Coefficients (Mel-logarithm cepstrum) of the speech segments, determined based on the integrated spectral envelopes (spectrum envelope; column 14, lines 3-22).

Regarding **claims 16, 41 and 66**, Itoh discloses a method, device and computer software product speech synthesis, comprising:

receiving an input speech signal (speech waveform) containing a set of speech segments (every phonemes segment; column 4, lines 59-67);

estimating spectral envelopes (spectrum envelope) of the input speech signal in a succession of time intervals during each of the speech segments (5ms; column 5, lines 19-35); and

integrating the spectral envelopes (spectrum envelope) over a plurality of window functions (window shifting) in a frequency domain so as to determine elements of feature vectors (parameter vector) corresponding to the speech segments (column 5, lines 19-35), but lacks feature domain concatenation.

Aso discloses a method for speech synthesis, comprising:

reconstructing an output speech signal by concatenating the feature vectors (figure 1 and figure 6) corresponding to a sequence of the speech segments (column 2, lines 55-67 with column 3, lines 15-22 and lines 53-67 and column 4, lines 1-8), obtaining a synthesized speech of higher quality.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Itoh's method, device and computer software produuct wherein it comprises feature domain concatenation, to obtain high-quality synthesized speech with low complexity in a more practical manner (column 2, lines 21-32).

Regarding claims **17, 42 and 67**, Itoh discloses a method, device and computer software product wherein receiving the input speech signal comprises dividing the input speech signal (partitioning) into the segments (phoneme segment) and determining segment information comprising respective phonetic identifiers (label indicating) of the segments, and wherein reconstructing the output speech signal comprises selecting the segments whose feature vectors are to be concatenated (clustering) responsive to the segment information determined with respect to the segments (column 4, line 59 – column 5, line 5.

Regarding **claims 18, 43 and 68**, Itoh discloses a method, device and computer software product wherein dividing the input speech signal into the segments comprises dividing the signal into lefemes (phoneme partitions), and wherein the phonetic identifiers comprise lefeme labels (labeling indicating combination of first, third phoneme; column 5, lines 1-17).

Regarding **claims 19, 44 and 69**, Itoh discloses a method, device and computer software product wherein determining the segment information further comprises finding respective segment parameters including one or more of a duration (duration), an energy level (power) and a pitch (pitch) of each of the segments, responsive to which parameters the segments are selected for use in reconstructing the output speech signal (column 6, lines 18-28 with column 9, lines 20-23).

Regarding **claims 20, 45 and 70**, Itoh discloses a method, device and computer software product wherein reconstructing the output speech signal (speech signal waveform) comprises modifying the feature vectors of the selected segments so as to

adjust the segment parameters of the segments in the output speech signal (modify the spectrum characteristic; column 7,lines 20-30).

Regarding **claims 21, 46 and 71**, Itoh discloses a method, device and computer software product and comprising determining respective degrees of voicing of the speech segments (degree of phoneme; column 9, lines 53-67), and incorporating the degrees of voicing as elements of the feature vectors for use in reconstructing the output speech signal (column 9, lines 53-67).

Regarding **claims 22, 47 and 72**, Itoh discloses a method, device and computer software product wherein concatenating the feature vectors comprises concatenating the vectors (column 6, lines 58-60) to form a series in a frequency domain (frequency domain; column 7, lines 30-32), and wherein reconstructing the output speech signal comprises computing a series of complex line spectra of the output signal from the series of feature vectors (parameter vector; column 5, lines 19-35), and transforming the complex line spectra to a time domain signal (time domain; column 8, lines 2-10).

5. **Claims 9, 34 and 59** are rejected under 35 U.S.C. 103(a) as being unpatentable over Itoh in view of Aso, as applied to claim 1 above, and in further view of Campbell et al. (U.S. Patent No. 6,366,883), hereinafter referenced as Campbell.

Regarding **claims 9, 34 and 59**, Itoh in view of Aso disclose a method, device and computer software product wherein selecting the sequences of feature vectors comprises:

selecting candidate segments from the inventory (Itoh; selects from each cluster a phoneme; column 5, lines 41-42), but lacks computing a cost function.

Campbell discloses a method, device and computer software product wherein selecting the sequences of feature vectors comprises:

computing a cost function for each of the candidate segments responsive to the phonetic (phoneme) and prosodic information (prosodic) and to the feature vectors of the candidate segments (phoneme candidates; column 5, lines 31-43); and

selecting the segments (searching phoneme sequences) so as to minimize the cost function (minimizes the cost; column 5,lines 31-43), for performing speech synthesis.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Itoh in combination with Aso's method, device and computer software product wherein the cost function is computed, for performing speech synthesis of any arbitrary sequence of phonemes by concatenation of speech segments of speech waveform signals extracted at synthesis time from a natural utterance (column 1, lines 12-16).

6. **Claims 11-12, 36-37 and 61-62** are rejected under 35 U.S.C. 103(a) as being unpatentable over Itoh in view of Aso, as applied to claim 1 above, and in further view of Mizuno et al. (U.S. Patent No. 6,334,106), hereinafter referenced as Mizuno.

Regarding **claims 11, 36 and 61**, Itoh in view of Aso disclose a method, device and computer software product, but lacks wherein the duration is shortened.

Mizuno discloses a method, device and computer software product wherein the prosodic information comprises respective durations of the segments to be incorporated in the output speech signal, and wherein adjusting the feature vectors (modifications of the dynamic range and envelope) comprises removing one or more of the feature vectors from the selected sequences so as to shorten the durations of one or more of the segments (shortening the duration; column 13, lines 23-30 and column 12, lines 20-22 with figure 7), to generate synthesized voices.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Itoh in combination with Aso's method, device and computer software product wherein the durations are shortened, to permit easy and fast synthesization of speech messages with desired prosodic features (column 1, lines 13-16).

Regarding **claims 12, 37 and 62**, Itoh in view of Aso disclose a method, device and computer software product, but lacks wherein the duration is lengthened.

Mizuno discloses a method, device and computer software product wherein the prosodic information comprises respective durations of the segments to be incorporated in the output speech signal, and wherein adjusting the feature vectors (modifications of the dynamic range and envelope) comprises adding one or more further feature vectors to the selected sequences so as to lengthen the durations of

one or more of the segments (lengthening the duration; column 13, lines 23-30 and column 12, lines 20-22 with figure 7), to generate synthesized voices.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Itoh in combination with Aso's method, device and computer software product wherein the durations are lengthened, to permit easy and fast synthesization of speech messages with desired prosodic features (column 1, lines 13-16).

7. **Claims 23-24, 48-49 and 73-74** are rejected under 35 U.S.C. 103(a) as being unpatentable over Itoh in view of Aso, as applied to claim 16 above, and in further view of Coorman et al. (U.S. Patent No. 6,665,641), hereinafter referenced as Coorman.

Regarding **claims 23, 48 and 73**, Itoh in view of Aso disclose a method, device and computer software product, but lacks wherein the window functions are non-zero only within different spectral windows.

Coorman discloses a method, device and computer software product wherein the window functions (limits) are non-zero only within different, respective spectral windows (non-zero outsides of limits) and have variable values over their respective windows (whole range; column 12, line 58 – column 13, line 6 and non-binary numeric; column 20, lines 36-37), and wherein integrating the spectral envelopes (spectral information; column 9,lines 45-47) comprises calculating products of the spectral envelopes with the window functions (optimized windowing), and calculating integrals

of the products over the respective windows of the window functions (column 20, lines 36-48), to maximize a similarity.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Itoh in combination with Aso's method, device and computer software product wherein the window functions are non-zero only within different spectral windows, for concatenation of the waveforms by maximizing a similarity measure between the windowed waveforms in a region near their adjacent edges (column 20, lines 38-48).

Regarding **claims 24, 49 and 74**, Itoh disclose a method, device and computer software product comprising applying a mathematical transformation to the integrals (equation 11) in order to determine the elements of the feature vectors (column 14, lines 3-22).

8. **Claims 25, 50 and 75** are rejected under 35 U.S.C. 103(a) as being unpatentable over Itoh in view of Aso and Coorman, as applied to claims 23, 48 and 73 above, in further view of Matsumoto (U.S. Patent No. 5,940,795).

Regarding **claims 25, 50 and 75**, Itoh in view of Aso and Coorman, as applied to claims 23, 48 and 73 above, discloses a method, device and computer software product wherein the frequency domain comprises a Mel frequency domain (Mel-logarithm), and wherein applying the mathematical transformation comprises applying log (logarithm) in order to determine Mel Frequency Cepstral Coefficients (Itoh; Mel-

logarithm cepstrum; column 14, lines 3-22) to be used as the elements of the feature vectors, but lacks discrete cosine transform operations.

Matsumoto discloses a method, device and computer software product comprising discrete cosine transformation operations (column 11, lines 50-64), for orthogonal transformation.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Itoh in combination with Aso and Coorman's method, device and computer software product comprising discrete cosine transformation operations, to conduct adaptive bit assignment for orthogonal transformation and to associated frequency analysis with audition of human beings, as taught by Matsumoto (column 11, lines 50-64).

Conclusion

9. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Jakieda R Jackson whose telephone number is 571.272.7619. The examiner can normally be reached on Monday through Friday from 7:30 a.m. to 5:00p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Wayne Young can be reached on 571.272.7582. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

JRJ
June 1, 2005

W. R. YOUNG
PRIMARY EXAMINER